AFRL-SA-WP-TR-2017-0014

Physical Profiling Performance of Air Force Primary Care Providers

Anthony P. Tvaryanas[1]; William P. Butler[1]; Brandon M. Greenwell[2]; Genny M. Maupin[1]; Valarie M. Schroeder[1]

[1]711th Human Performance Wing; [2]InfoSciTex, a DCS Company

August 2017

Final Report
for September 2016 to January 2017

Air Force Research Laboratory
711th Human Performance Wing
U.S. Air Force School of Aerospace Medicine
Aeromedical Research Department
2510 Fifth St., Bldg. 840
Wright-Patterson AFB, OH 45433-7913

NOTICE AND SIGNATURE PAGE

AFRL-SA-WP-TR-2017-0014 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNATURE//	//SIGNATURE//
MS. LINDA ARMSTRONG Chief, Airman Integration Res Support Div	DR. RICHARD A. HERSACK Chair, Aeromedical Research Department

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
9 Aug 2017	Final Technical Report	September 2016 – January 2017

4. TITLE AND SUBTITLE

Physical Profiling Performance of Air Force Primary Care Providers

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Anthony P. Tvaryanas, William P. Butler, Brandon M. Greenwell, Genny M. Maupin, Valarie M. Schroeder

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

USAF School of Aerospace Medicine
Aeromedical Research Dept/FHS
2510 Fifth St., Bldg. 840
Wright-Patterson AFB, OH 45433-7913

8. PERFORMING ORGANIZATION REPORT NUMBER

AFRL-SA-WP-TR-2017-0014

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited.

13. SUPPLEMENTARY NOTES

Cleared, 88PA, Case # 2017-4145, 28 Aug 2017.

14. ABSTRACT

The purpose of this study was to determine the test performance of primary care providers (PCPs) in screening for occupational limitations and ascertain if predictors existed to augment PCP screening. This study was a cross-sectional, retrospective medical records review of active duty U.S. Air Force (AF) members receiving care in an AF medical treatment facility (MTF) between October 31, 2013, and September 30, 2014, who had at least one encounter with their primary care team. An independent medical standards subject matter expert (SME) reviewed encounters in the electronic health record and determined whether duty, fitness, and/or mobility restrictions were indicated. PCP dispositions were obtained from archival data as were service member age, sex, job category, and years of military service. Encounter-related archival data included diagnosis; encounter type (Periodic Health Assessment [PHA] versus non-PHA); associated laboratory, radiology, and specialty consult orders; and provider type. Nonparametric and parametric models were used to identify variables associated with SME-identified occupational limitations and PCP-SME disagreement on occupational limitations. The proportion of participants identified as having occupational limitations significantly differed between the PCP and the SME. PCP screening test performance differed by encounter type and was generally better for PHA encounters. Overall, the models of SME-identified occupational limitations were comprised of a small number of predictors, and the performance of the models was generally fair. The models of PCP-SME disagreement did not find any strong and consistent associations across restriction types, and the overall performance of the models was only poor to fair. PCPs appear to adjust their screening behaviors based on encounter type, and when factoring in differences in the prevalence of service members with occupational limitations, profiling quality assurance activities may yield greater return on investment if focused on reviewing non-PHA encounters. The study modeling results do not suggest that a computational solution implemented within an electronic health record system will be forthcoming in the near term that can either augment or alleviate the PCP in screening for occupational limitations. Meanwhile, continued training and quality assurance activities with feedback to PCPs to enable continuous learning are the most tractable near-term approach for improving profiling system performance.

15. SUBJECT TERMS

Medical treatment facility, occupational limitations, periodic health assessments, screening, statistical modeling

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Ms. Linda Armstrong
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	SAR	23	

TABLE OF CONTENTS

Section	Page
LIST OF FIGURES	ii
LIST OF TABLES	ii
1.0 EXECUTIVE SUMMARY	1
2.0 INTRODUCTION	2
3.0 METHODS	2
3.1 Study Design and Participants	2
3.2 Procedure and Measurements	3
3.3 Statistical Analysis	3
4.0 RESULTS	5
4.1 Descriptive Statistics	5
4.2 Contingency Table Analysis	5
4.3 Prediction Models	5
5.0 DISCUSSION	14
6.0 REFERENCES	16
LIST OF ABBREVIATIONS AND ACRONYMS	17

LIST OF FIGURES

Page

Figure 1. Procedure for contingency table analysis of screening for occupational limitations 4

Figure 2. Relative influence of each predictor based on XGBoost (zero importance indicates the variable made no contributions to predicting the response) 8

LIST OF TABLES

Page

Table 1. Study Variables and Associated Summary Descriptive Statistics 6

Table 2. Contingency Table Analyses for PCP vs. SME Decision Making on 7

Table 3. Primary Care Provider Performance for Screening for Occupational Limitations 7

Table 4. AUROC Curve Statistics for XGBoost and Logistic Regression Models 7

Table 5. Logistic Regression Model Results for SME-Identified Duty Restrictions 10

Table 6. Logistic Regression Model Results for SME-Identified Mobility Restrictions 11

Table 7. Logistic Regression Model Results for SME-Identified Fitness Restrictions 12

Table 8. Logistic Regression Model Results for PCP-SME Disagreement on Duty Restrictions 13

Table 9. Logistic Regression Model Results for PCP-SME Disagreement on Mobility Restrictions 14

Table 10. Logistic Regression Model Results for PCP-SME Disagreement on Fitness Restrictions 15

1.0 EXECUTIVE SUMMARY

Primary care providers (PCPs) are responsible for performing an occupational disposition at the end of every healthcare encounter. The occupational disposition addresses whether a service member is capable of unrestricted activity or requires duty, fitness, and/or mobility restrictions. The purpose of this study was to determine the test performance of PCPs in screening for occupational limitations and ascertain if predictors existed to augment PCP screening.

This study was a cross-sectional, retrospective medical records review of active duty U.S. Air Force (AF) members receiving care in an AF medical treatment facility (MTF) between October 31, 2013, and September 30, 2014, who had at least one encounter with their primary care team. The study used a stratified random sample on encounters from all AF MTFs. An independent medical standards subject matter expert (SME) reviewed encounters in the electronic health record and determined whether duty, fitness, and/or mobility restrictions were indicated. PCP dispositions were obtained from archival data as were service member age, sex, job category, and years of military service. Encounter-related archival data included diagnosis; encounter type (Periodic Health Assessment [PHA] versus non-PHA); associated laboratory, radiology, and specialty consult orders; and provider type. PCP sensitivity, specificity, and positive and negative predictive values were calculated based on 2 x 2 contingency table analyses. Nonparametric and parametric models were used to identify variables associated with SME-identified occupational limitations and PCP-SME disagreement on occupational limitations.

The proportion of participants identified as having occupational limitations significantly differed between the PCP and the SME for all three restriction types (i.e., duty, mobility, and fitness). Overall PCP sensitivity, specificity, and positive and negative predictive values were 84.28%, 84.99%, 76.60%, and 90.27%, respectively. PCP screening test performance differed by encounter type and was generally better for PHA encounters. Overall, the models of SME-identified occupational limitations identified a small number of predictors, and the performance of the models was generally fair. Only the diagnosis conditions of injury and poisoning and musculoskeletal and connective tissue diseases and male sex were associated with occupational limitations. The models of PCP-SME disagreement did not find any strong and consistent associations across restriction types, and the overall performance of the models was only poor to fair.

PCPs appear to adjust their screening behaviors based on encounter type, and when factoring in differences in the prevalence of service members with occupational limitations, profiling quality assurance activities may yield greater return on investment if focused on reviewing non-PHA encounters. The study modeling results do not suggest that a computational solution implemented within an electronic health record system will be forthcoming in the near term that can either augment or alleviate the PCP in screening for occupational limitations. Meanwhile, continued training and quality assurance activities with feedback to PCPs to enable continuous learning are the most tractable near-term approach for improving profiling system performance.

2.0 INTRODUCTION

Primary care providers (PCPs) are responsible for performing an occupational disposition at the end of every healthcare encounter. The occupational disposition addresses whether a service member is capable of unrestricted activity or requires duty, fitness, and/or mobility restrictions. When restrictions are identified, they are documented and communicated to supervisors and commanders using a physical profile (Air Force [AF] Form 469). Presently, PCPs do not receive standardized training on physical profiling.

An analysis of PCP physical profiles generated at one medical treatment facility (MTF) found poor agreement between the PCPs and independent subject matter expert (SME) reviewers (kappa = 0.152, 95% confidence interval [CI]: 0.047-0.258). Additionally, the analysis found that PCPs' recommendation of a specialty referral at an encounter was strongly associated with an independent reviewer judging that a service member required a physical profile (odds ratio [OR] = 8.889, 95% CI: 2.226-41.209).[1] Together, these observations suggest that PCPs, at one MTF at least, do not perform well at the task of physical profiling, and there may be potential indicators to help identify service members who warrant more scrutiny in terms of requiring a physical profile.

The performance of the physical profiling system is critical to providing an accurate measure of service member availability to line commanders. Underprofiling puts individual service members at risk for aggravation of underlying medical conditions and/or deployment to locations where there are inadequate healthcare resources to meet their medical needs. Additionally, service members who are unable to adequately perform their duties because of physical limitations potentially put unit mission accomplishment at risk. The purpose of this study was to ascertain whether the abovementioned observations generalized beyond the one MTF and to explore whether there were other potential indicators for service members who are at increased risk for needing a physical profile.

3.0 METHODS

3.1 Study Design and Participants

This study was conducted under a human-use protocol approved by the 711th Human Performance Wing Institutional Review Board. A waiver of informed consent of participants was granted due to the impracticality of obtaining written consent from each participant in the study population. This study was a cross-sectional, retrospective medical records review of active duty AF members receiving care in an AF MTF between October 31, 2013, and September 30, 2014, who had at least one encounter with their primary care team. Participants were excluded from the study if they received primary care in an MTF of a sister service. The study sample was assembled by selecting a stratified random sample from all AF MTFs, taking care to ensure that each stratified sample was proportional in size to the total number of military personnel enrolled at the corresponding MTF.

[1] Unpublished data from an MTF-level process improvement project.

3.2 Procedure and Measurements

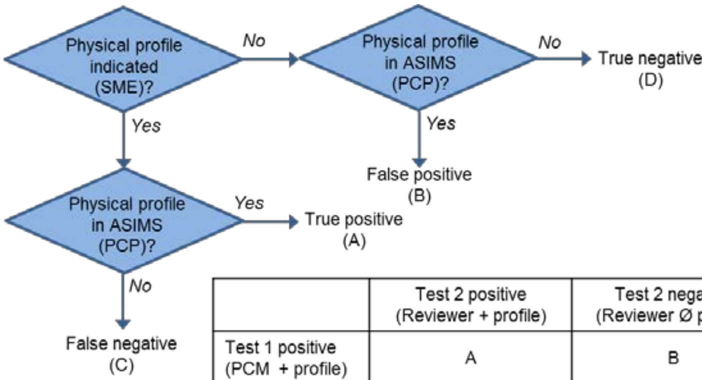
Data from the Military Health System Data Mart were used to generate the stratified random sample of participant primary care team encounters. This list was provided to an independent medical standards SME who reviewed the corresponding clinical encounter notes in the electronic health record ─ that is, the Armed Forces Health Longitudinal Technology Application ─ to determine whether duty, fitness, and/or mobility restrictions were indicated. The SME also determined the encounter type (Periodic Health Assessment [PHA] or non-PHA) and recorded any associated radiology or specialty referral orders.

For each participant encounter, data were extracted from the Military Health System Data Mart on provider identification; International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnosis codes; and the number of prior encounters within the preceding 60 days and 365 days. Data were also obtained from the Aeromedical Services Information Management System on participants' existing physical profiles to include restriction types (duty, fitness, and/or mobility) and associated ICD-9-CM codes. The Air Force Personnel Center database was used to obtain participants' age, sex, primary Air Force Specialty Code (AFSC), and years of military service as well as determine providers' medical specialty code. Provider medical specialty codes included primary care nurse practitioner, physician assistant, family practice physician, independent duty medical technician (IDMT), flight surgeon, or other.

The various datasets were merged using participant Social Security number and date of birth, the latter to ensure data were related to a service member rather than a dependent. Social Security number and date of birth were then removed from the composite study dataset. ICD-9-CM codes were recoded using the Healthcare Cost and Utilization Project Clinical Classification Software for ICD-9-CM, which aids analysts in collapsing diagnostic data from over 14,000 diagnosis codes into clinically meaningful categories [1]. Diagnosis categories used in this study included the following: neoplasms; mental disorders; injury and poisoning; infectious and parasitic diseases; endocrine, nutritional/metabolic diseases, and immunity disorders; skin and subcutaneous tissue diseases; respiratory diseases; nervous system/sense organ diseases; musculoskeletal and connective tissue diseases; genitourinary system diseases; digestive system diseases; circulatory system diseases; ill-defined conditions; and supplementary factors influencing health status.

3.3 Statistical Analysis

Based on the procedure depicted in Figure 1, 2 x 2 contingency tables were created where the PCPs were considered a screening test for occupational limitations ("Test 1") and the independent SME was considered the "gold standard" test for occupational limitations ("Test 2"). The contingency table analysis was stratified by encounter type (PHA versus non-PHA). Contingency tables were created for duty, fitness, and mobility restrictions. Sensitivity, specificity, positive predictive value, and negative predictive value were calculated for each contingency table.



	Test 2 positive (Reviewer + profile)	Test 2 negative (Reviewer Ø profile)	Row totals
Test 1 positive (PCM + profile)	A	B	A+B
Test 1 negative (PCM Ø profile)	C	D	C+D
Column totals	A+C	B+D	A+B+C+D

Figure 1. Procedure for contingency table analysis of screening for occupational limitations.

Nonparametric and parametric models were used to identify predictors of participants requiring a physical profile (duty, mobility, and fitness restrictions) as judged by the independent SME (cells A and C) and predictors of PCP-SME disagreement (cells B and C). Potential predictors included encounter diagnosis categories; encounter type (reduced to PHA versus non-PHA); lab (reduced to yes or no), radiology, (reduced to none, computed tomography scan, magnetic resonance imaging (MRI), and other X-rays), and specialty referral (reduced to yes or no) orders; number of prior encounters within the preceding 60 days and 365 days; participant age, sex (male or female), job category based on primary AFSC (reduced to acquisition, logistics and maintenance, medical, operations, professional, reporting identifiers, special duty identifiers, special investigations, support, and not available), and years of military service; and provider type (reduced to aerospace medicine physician [inclusive of 48A and 48G AFSCs], family practice physician, IDMT, nurse practitioner, other provider, and physician assistant). Prior to analysis, all high-level categorical variables (i.e., encounter diagnosis categories, AFSC, and provider specialty) were one-hot encoded; that is, a separate dummy variable was created for every level of each variable. This procedure yielded a total of 40 predictor variables that were used for modeling. All statistical analyses were accomplished using R version 3.3.0 [2]. Statistical significance was a priori defined at the 0.05 level.

The final dataset comprised 812 unique participant encounters with only one observation per participant. The dataset was split into two samples at random: 500 observations for exploratory learning and variable selection and 312 observations for validation and inference. Since the recommended minimum number of observations per variable suggested for logistic regression is 10 [3], the sample size for validation and inference was large enough to obtain reasonably stable regression coefficients. Nonparametric methods were used for exploratory analysis and parametric methods were used for model building and statistical inference given the greater ease of interpretation of the latter. Separating variable selection and model building ensured that the reported standard errors (SEs), and hence the corresponding *p*-values, were valid. Additionally, the use of a smaller dataset for model building helped control for the effect of sample size on *p*-values.

Extreme gradient boosting [4] (denoted henceforth as XGBoost) is a popular machine learning library based on stochastic gradient boosting [5,6]. XGBoost was used for exploratory analysis on the learning sample. XGBoost has the ability to ignore irrelevant predictors and rank

the import variables in terms of their relative influence in separating the response categories. Larger relative influence scores suggest greater importance in terms of predicting the response. These scores range from 0, indicating an irrelevant predictor, to 1, indicating the most important predictor. For each response variable, the corresponding XGBoost model was trained to maximize the area under the receiver operating characteristic (AUROC) curve. All predictors with nonzero influence for a particular response were included in a logistic regression model to estimate an OR and to assess for statistical significance at the 0.05 level.

4.0 RESULTS

4.1 Descriptive Statistics

Table 1 provides the study variables and associated summary descriptive statistics for the overall study sample.

4.2 Contingency Table Analysis

Table 2 provides a summative view of the multiple 2 x 2 contingency table analyses of PCP versus SME decision making regarding occupational limitations stratified by encounter type and restriction type. Based on the chi-square statistic, the proportion of participants identified as having occupational limitations significantly differed between the PCP and the SME for all three restriction types (i.e., duty, mobility, and fitness) regardless of encounter type (i.e., PHA versus non-PHA). Table 3 summarizes the PCP performance for screening for occupational limitations as compared to the SME “gold standard.” Out of the 243 participants who were assigned a fitness restriction by the PCP, 136 participants received a concurrent duty restriction (a proportion of 55.97%) and 116 participants received a concurrent mobility restriction (a proportion of 47.74%). Similarly, of the 240 participants who were assigned a fitness restriction by the SME, 147 participants received a concurrent duty restriction (a proportion of 61.25%) and 156 participants received a concurrent mobility restriction (a proportion of 65.00%).

4.3 Prediction Models

Six XGBoost models and six logistic regression models were generated to evaluate the two responses (i.e., SME-identified restrictions and PCP-SME disagreement) for each restriction type (i.e., duty, mobility, and fitness). The AUROC curve statistics for the XGBoost models and corresponding logistic regression models (after XGBoost variable selection) are shown in Table 4. Considering the AUROC curve statistics for the XGBoost models, which provide a more honest estimate, predictive accuracy of the models was generally poor to fair, with the exception of the model of SME-identified fitness restrictions, where accuracy was good.

Table 1. Study Variables and Associated Summary Descriptive Statistics

Variable	Descriptive Statistic
Diagnoses, no. (%):	
Circulatory system diseases	9 (1.11)
Digestive system diseases	13 (1.60)
Endocrine, nutritional/metabolic diseases and immunity disorders	9 (1.11)
Genitourinary system diseases	8 (0.99)
Ill-defined conditions	41 (5.05)
Infectious and parasitic diseases	16 (1.97)
Injury and poisoning	33 (4.06)
Mental disorders	11 (1.35)
Musculoskeletal and connective tissue diseases	126 (15.52)
Neoplasms	3 (0.37)
Nervous system/sense organ diseases	23 (2.83)
Respiratory diseases	48 (5.91)
Skin and subcutaneous tissue diseases	21 (2.59)
Supplementary factors influencing health status	451 (55.54)
Encounter-related orders, no. (%):	
Computed tomography scan	3 (0.37)
Labs	117 (14.41)
MRI	14 (1.72)
Other X-ray	64 (7.88)
Specialty referrals	153 (18.84)
Encounter type, no. (%)	
Non-PHA	446 (54.93)
PHA	366 (45.07)
Healthcare utilization, median (IQR):	
Encounters in last 60 days	0 (1)
Encounters in last 365 days	3 (4)
Job category, no. (%):	
Acquisition	27 (3.33)
Logistics and maintenance	245 (31.21)
Medical	102 (8.51)
Operations	173 (13.84)
Professional	11 (1.02)
Reporting identifiers	11 (0.90)
Special duty identifiers	12 (0.64)
Special investigation	3 (0.16)
Support	222 (11.99)
Not available	6 (0.37)
Participant characteristics:	
Age, yr, median (IQR)	29.61 (11.64)
Sex, no. (%):	
Male	665 (80.67)
Female	152 (18.72)
Missing	5 (0.62)
Years of service, median (IQR)	7.99 (11.8)
Provider type, no. (%):	
Aerospace medicine physician	98 (12.07)
Family practice physician	223 (27.46)
IDMT	96 (11.82)
Nurse practitioner	100 (12.32)
Other provider	66 (8.13)
Physician assistant	229 (28.20)

IQR = interquartile range.

Table 2. Contingency Table Analyses for PCP vs. SME Decision Making on Occupational Limitations

Encounter Type	Restriction Type	Restriction by PCP?	Restriction by SME? Yes	No	Chi-Square	*p*-value
PHA	Duty	Yes	36	5	219.090	<0.0001
		No	15	310		
	Mobility	Yes	38	8	192.602	<0.0001
		No	16	304		
	Fitness	Yes	42	10	187.349	<0.0001
		No	17	297		
Non-PHA	Duty	Yes	118	40	174.515	<0.0001
		No	36	252		
	Mobility	Yes	56	66	56.085	<0.0001
		No	42	282		
	Fitness	Yes	136	52	129.575	<0.0001
		No	48	210		

Table 3. Primary Care Provider Performance for Screening for Occupational Limitations

Encounter Type	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
PHA (95% CI)	93.55% (84.30%, 98.21%)	90.79% (86.96%, 93.79%)	67.44% (59.13%, 74.78%)	98.57% (96.39%, 99.44%)
Non-PHA (95% CI)	81.86% (76.35%, 86.55%	76.56% (70.22%, 82.12%)	79.84% (75.47%, 83.59%)	78.82% (73.76%, 83.13%)

Table 4. AUROC Curve Statistics for XGBoost and Logistic Regression Models

Response	Restriction Type	Model: XGBoost	Model: Logistic Regression
SME-identified restrictions	Duty	0.737	0.915
	Mobility	0.778	0.879
	Fitness	0.822	0.898
PCP-SME disagreement	Duty	0.700	0.983
	Mobility	0.669	0.822
	Fitness	0.703	0.903

The relative influence for each predictor from all six XGBoost models is displayed in Figure 2. By inspection, years of service, sex, respiratory diagnoses, other healthcare services diagnoses, encounter type, encounters within the last 60 days and 365 days, and age were the more important variables in explaining SME-identified duty restrictions. Years of service, sex, respiratory disease diagnoses, musculoskeletal and connective tissue disease diagnoses, family practice physician provider type, encounters within the last 365 days, and age were the more important variables in explaining SME-identified mobility restrictions. Years of service, musculoskeletal and connective tissue disease diagnoses, injury and poisoning diagnoses, and encounters within the last 365 days were the more important variables in explaining SME-identified mobility restrictions.

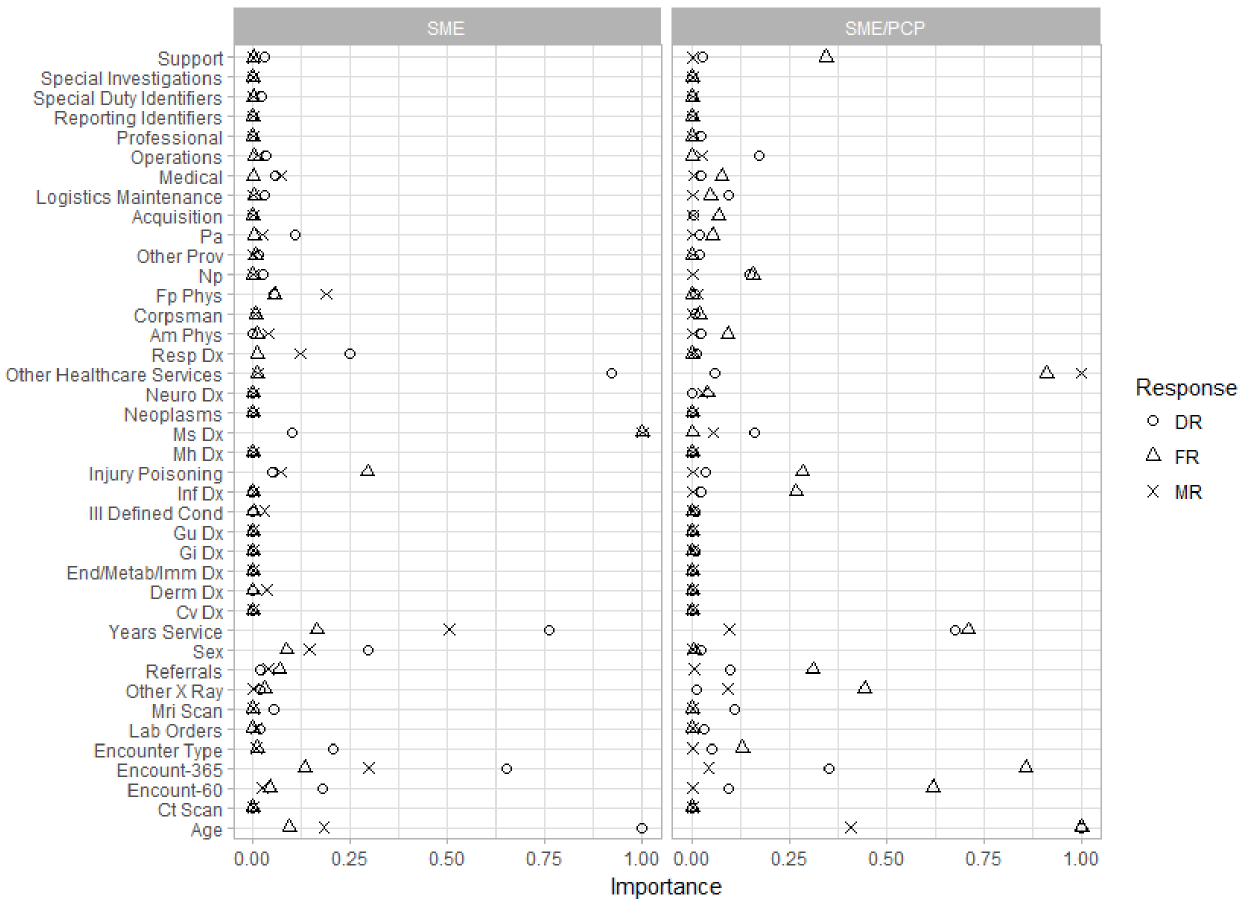


Figure 2. Relative influence of each predictor based on XGBoost (zero importance indicates the variable made no contributions to predicting the response).

Similarly, years of service, operations job category, nurse practitioner provider type, musculoskeletal and connective tissue disease diagnoses, encounters within the last 365 days, and age were the more important variables in explaining PCP-SME disagreement on duty restrictions. Only other healthcare services diagnoses and nurse practitioner provider type were important variables in explaining PCP-SME disagreement on mobility restrictions. Years of service, support job category, specialty referrals, other X-ray orders, supplementary factors influencing health status diagnoses, injury and poisoning diagnoses, infectious and parasitic disease diagnoses, encounter type, and encounters within the last 60 and 365 days were the most important variables in explaining PCP-SME disagreement on fitness restrictions.

The estimated regression coefficients, including approximate SEs and *p*-values, for the logistic regression models for SME-identified duty, mobility, and fitness restrictions are displayed in Tables 5, 6, and 7, respectively. Injury and poisoning (OR: 12.231; 95% CI: 3.198, 63.137) and musculoskeletal and connective tissue disease diagnoses (OR: 3.140; 95% CI: 1.227, 8.296) were associated with an increased likelihood for duty restrictions, while supplementary factors influencing health status diagnoses (OR: 0.208; 95% CI: 0.048, 0.715) and IDMT provider type (OR: 0.143; 95% CI: 0.016, 0.828) were associated with a decreased likelihood for duty restrictions. Injury and poisoning (OR: 11.963; 95% CI: 3.076, 53.179) and musculoskeletal and connective tissue disease diagnoses (OR: 4.330; 95% CI: 1.537, 12.888) were associated with an increased likelihood for mobility restrictions, while male sex (OR: 0.387; 95% CI: 0.168, 0.898) was associated with a decreased likelihood for mobility restrictions. Likewise, injury and poisoning (OR: 45.924; 95% CI: 9.622, 351.986) and musculoskeletal and connective tissue disease diagnoses (OR: 15.116; 95% CI: 5.029, 50.425) were associated with an increased likelihood for fitness restrictions, while male sex (OR: 0.383; 95% CI: 0.161, 0.918) was associated with a decreased likelihood for fitness restrictions.

The estimated regression coefficients, including approximate SEs and *p*-values, for the logistic regression models for PCP-SME disagreement on duty, mobility, and fitness restrictions are displayed in Tables 8, 9, and 10, respectively. PHA encounter type (OR: 7.610; 95% CI: 1.590, 43.598) was associated with an increased likelihood for disagreement, while injury and poisoning (OR: 0.163; 95% CI: 0.028, 0.904) and musculoskeletal and connective tissue disease diagnoses (OR: 0.209; 95% CI: 0.046, 0.795) were associated with a decreased likelihood for disagreement on duty restrictions. Family practice physician provider type (OR: 0.440; 95% CI: 0.191, 0.989) was associated with a decreased likelihood for disagreement on mobility restrictions. PHA encounter type (OR: 3.775; 95% CI: 1.058, 13.572) was associated with an increased likelihood for disagreement, while other X-ray orders (OR: 0.279; 95% CI: 0.092, 0.863) and encounters in the last 365 days (OR: 0.808; 95% CI: 0.709, 0.913) were associated with a decreased likelihood for disagreement on fitness restrictions.

Table 5. Logistic Regression Model Results for SME-Identified Duty Restrictions

Variable	*B*	SE (*B*)	OR	*p*-value
Diagnoses:				
Injury and poisoning	2.504	0.742	12.231	0.001
Musculoskeletal and connective tissue diseases	1.144	0.485	3.140	0.018
Respiratory diseases	0.539	0.599	1.714	0.368
Supplementary factors influencing health status	-1.571	0.678	0.208	0.021
Encounter-related orders:				
Labs (yes)	0.132	0.441	1.141	0.765
MRI (yes)	2.145	1.272	8.539	0.092
Other X-ray (yes)	-1.014	0.616	0.363	0.100
Specialty referrals (yes)	-0.059	0.391	0.943	0.880
Encounter type (PHA)	0.308	0.682	1.361	0.651
Healthcare utilization:				
Encounters in last 60 days	0.120	0.220	1.128	0.586
Encounters in last 365 days	-0.042	0.056	0.959	0.456
Job category:				
Logistics and maintenance	0.360	0.705	1.434	0.609
Medical	0.422	0.826	1.526	0.609
Operations	0.951	0.764	2.588	0.213
Special duty identifiers	1.712	1.573	5.540	0.276
Support	0.640	0.719	1.896	0.373
Participant characteristics:				
Age	-0.080	0.062	0.923	0.200
Sex (male)	-0.455	0.423	0.634	0.282
Years of service	0.046	0.065	1.047	0.474
Provider type:				
Family practice physician	-0.352	0.630	0.703	0.576
IDMT	-1.948	0.965	0.143	0.044
Nurse practitioner	-0.562	0.746	0.570	0.451
Other provider	-0.659	0.727	0.517	0.364
Physician assistant	-0.570	0.639	0.566	0.372

Table 6. Logistic Regression Model Results for SME-Identified Mobility Restrictions

Variable	*B*	SE (*B*)	OR	*p*-value
Diagnoses:				
Ill-defined conditions	-1.034	0.860	0.356	0.229
Infectious and parasitic diseases	-14.734	2399.545	0.000	0.995
Injury and poisoning	2.482	0.720	11.963	0.001
Musculoskeletal and connective tissue diseases	1.466	0.540	4.330	0.007
Respiratory diseases	-1.270	1.120	0.281	0.257
Skin and subcutaneous tissue diseases	-15.708	953.048	<0.001	0.987
Supplementary factors influencing health status	-0.680	0.663	0.506	0.305
Encounter-related orders:				
Labs (yes)	0.173	0.473	1.189	0.715
Other X-ray (yes)	0.343	0.600	1.409	0.568
Specialty referrals (yes)	0.506	0.385	1.658	0.189
Encounter type (PHA)	0.019	0.647	1.019	0.977
Healthcare utilization:				
Encounters in last 60 days	0.054	0.213	1.056	0.800
Encounters in last 365 days	0.026	0.056	1.026	0.647
Job category:				
Medical	0.477	0.640	1.611	0.456
Operations	0.801	0.491	2.228	0.103
Support	0.368	0.416	1.445	0.375
Participant characteristics:				
Age	-0.048	0.058	0.953	0.406
Sex (male)	-0.950	0.425	0.387	0.026
Years of service	0.076	0.062	1.079	0.215
Provider type:				
Aerospace medicine physician	-0.484	0.748	0.616	0.517
Family practice physician	0.741	0.469	2.097	0.114
IDMT	-0.858	0.884	0.424	0.332
Physician assistant	-0.175	0.474	0.840	0.712

Table 7. Logistic Regression Model Results for SME-Identified Fitness Restrictions

Variable	*B*	SE (*B*)	OR	*p*-value
Diagnoses:				
Ill-defined conditions	-1.173	0.922	0.309	0.203
Injury and poisoning	3.827	0.886	45.924	<0.001
Musculoskeletal and connective tissue diseases	2.716	0.585	15.116	<0.001
Respiratory diseases	-0.029	0.778	0.972	0.970
Supplementary factors influencing health status	-0.079	0.612	0.924	0.897
Encounter-related orders:				
Other X-ray (yes)	1.130	0.705	3.096	0.109
Specialty referrals (yes)	-0.174	0.429	0.840	0.685
Encounter type (PHA)	-0.587	0.585	0.556	0.315
Healthcare utilization:				
Encounters in last 60 days	0.421	0.218	1.523	0.054
Encounters in last 365 days	-0.119	0.067	0.888	0.077
Job category:				
Logistics and maintenance	1.029	0.848	2.798	0.225
Medical	0.525	0.953	1.690	0.582
Operations	0.760	0.894	2.139	0.395
Special duty identifiers	-13.157	805.048	0.000	0.987
Support	0.994	0.863	2.703	0.249
Participant characteristics:				
Age	-0.038	0.063	0.963	0.549
Sex (male)	-0.959	0.442	0.383	0.030
Years of service	0.105	0.066	1.110	0.111
Provider type:				
Aerospace medicine physician	0.561	0.810	1.752	0.489
Family practice physician	0.684	0.594	1.981	0.250
IDMT	0.086	0.840	1.089	0.919
Other provider	0.408	0.759	1.503	0.591
Physician assistant	0.244	0.593	1.276	0.681

Table 8. Logistic Regression Model Results for PCP-SME Disagreement on Duty Restrictions

Variable	*B*	SE (*B*)	OR	*p*-value
Diagnoses:				
Digestive system diseases	-1.783	1.086	0.168	0.101
Ill-defined conditions	0.135	1.201	1.145	0.910
Injury and poisoning	-1.813	0.867	0.163	0.037
Musculoskeletal and connective tissue diseases	-1.565	0.713	0.209	0.028
Respiratory diseases	-0.971	1.007	0.379	0.335
Supplementary factors influencing health status	-1.069	0.772	0.343	0.166
Encounter-related orders:				
Labs (yes)	0.836	0.700	2.307	0.232
MRI (yes)	-0.924	1.076	0.397	0.390
Other X-ray (yes)	0.366	0.776	1.441	0.638
Specialty referrals (yes)	0.129	0.492	1.138	0.792
Encounter type (PHA)	2.029	0.823	7.610	0.014
Healthcare utilization:				
Encounters in last 60 days	-0.060	0.253	0.942	0.813
Encounters in last 365 days	-0.078	0.067	0.925	0.245
Job category:				
Acquisition	-0.009	1.839	0.991	0.996
Logistics and maintenance	-0.354	1.415	0.702	0.802
Medical	-1.027	1.491	0.358	0.491
Operations	-0.043	1.452	0.958	0.976
Support	-0.124	1.417	0.884	0.930
Participant characteristics:				
Age	-0.059	0.075	0.943	0.430
Sex (male)	0.354	0.594	1.425	0.551
Years of service	0.009	0.079	1.009	0.914
Provider type:				
Aerospace medicine physician	-14.492	989.555	<0.001	0.988
Family practice physician	-16.020	989.554	<0.001	0.987
Nurse practitioner	-15.605	989.554	<0.001	0.987
Other provider	-16.424	989.554	<0.001	0.987
Physician assistant	-15.873	989.554	<0.001	0.987

Table 9. Logistic Regression Model Results for PCP-SME Disagreement on Mobility Restrictions

Variable	*B*	SE (*B*)	OR	*p*-value
Diagnoses:				
Musculoskeletal and connective tissue diseases	-0.752	0.468	0.471	0.108
Nervous system/sense organ diseases	-0.089	0.751	0.915	0.906
Supplementary factors influencing health status	0.618	0.394	1.856	0.117
Encounter-related orders:				
Other X-ray (yes)	-0.757	0.553	0.469	0.171
Specialty referrals (yes)	0.124	0.394	1.132	0.752
Healthcare utilization:				
Encounters in last 365 days	-0.060	0.046	0.942	0.190
Job category:				
Medical	-0.512	0.533	0.599	0.337
Operations	-0.582	0.423	0.559	0.169
Participant characteristics:				
Age	-0.035	0.055	0.965	0.523
Years of service	0.014	0.059	1.014	0.810
Provider type:				
Family practice physician	-0.822	0.416	0.440	0.048
Physician assistant	-0.361	0.433	0.697	0.404

5.0 DISCUSSION

The purpose of this study was to determine the test performance of PCPs in screening for occupational limitations and ascertain if predictors existed to augment PCP screening. PCPs were observed to have a 0.84 probability of initiating a profile when a service member required an occupational limitation. PCPs were also observed to have a 0.85 probability of not initiating a profile when a service member did not require an occupational limitation. There was a noteworthy difference in PCP screening performance between PHA and non-PHA encounters, with both superior sensitivity and specificity observed during the PHA encounters. This finding suggests that PCPs adjust their behaviors based on encounter type.

While sensitivity and specificity are characteristics of the PCPs themselves, predictive values are affected by the prevalence of occupational limitations in the population being screened. Again, noticeable differences were observed between PHA and non-PHA encounters. However, the prevalence of service members with occupational limitations was 16.94% for PHA encounters and 53.14% for non-PHA encounters. When the observed sensitivities and specificities are applied to a hypothetical population with a 20% prevalence of occupational limitations, the positive predictive value for the PCP in PHA versus non-PHA encounters would be 71.75% and 46.61%, respectively. Likewise, the negative predictive value for the PCP in PHA versus non-PHA encounters would be 98.25% versus 94.40%, respectively. Since the PCP false negative or “miss” rate is 6.45% versus 18.14% for PHA versus non-PHA encounters, respectively, profiling quality assurance activities may yield greater return on investment if focused on reviewing non-PHA encounters.

Table 10. Logistic Regression Model Results for PCP-SME Disagreement on Fitness Restrictions

Variable	*B*	SE (*B*)	OR	*p*-value
Diagnoses:				
Infectious and parasitic diseases	-17.396	882.744	0.000	0.984
Injury and poisoning	-0.124	0.742	0.884	0.868
Nervous system/sense organ diseases	0.882	0.971	2.415	0.364
Supplementary factors influencing health status	-0.166	0.570	0.847	0.771
Encounter-related orders:				
Other X-ray (yes)	-1.276	0.565	0.279	0.024
Specialty referrals (yes)	0.347	0.457	1.415	0.447
Encounter type (PHA)	1.328	0.644	3.775	0.039
Healthcare utilization:				
Encounters in last 60 days	0.056	0.230	1.058	0.806
Encounters in last 365 days	-0.213	0.064	0.808	0.001
Job category:				
Acquisition	0.515	1.265	1.674	0.684
Logistics and maintenance	0.188	0.564	1.207	0.739
Medical	0.807	0.944	2.241	0.393
Support	-0.362	0.571	0.696	0.525
Participant characteristics:				
Age	0.023	0.070	1.024	0.737
Sex (male)	0.198	0.489	1.219	0.685
Years of service	-0.043	0.072	0.958	0.552
Provider type:				
Aerospace medicine physician	-0.235	0.718	0.790	0.743
IDMT	-0.636	0.734	0.529	0.386
Nurse practitioner	0.501	0.705	1.650	0.477
Physician assistant	-0.005	0.438	0.995	0.990

The modeling of SME-determined duty, mobility, and fitness restrictions sought to identify data-driven predictors that could be used to enhance PCP screening performance. Overall, the models identified a small number of predictors, and the performance of the models was generally fair. Only the diagnosis conditions of injury and poisoning and musculoskeletal and connective tissue diseases and male sex were associated with occupational limitations. These results do not suggest that a computational solution implemented within an electronic health record system will be forthcoming in the near term that can either augment or alleviate the PCP in screening for occupational limitations. An additional interesting observation was the absence of a generalized association with provider type, as aerospace medicine providers receive dedicated profile training and are assumed to be SMEs within the MTFs.

Modeling of PCP-SME disagreements on duty, mobility, and fitness restrictions sought to identify data-driven predictors for patients that would warrant a second look by a provider with expertise in identifying occupational limitations. Unfortunately, no strong and consistent associations were observed across restriction types, and the overall performance of the models was only poor to fair. Accordingly, profiling quality assurance activities needs to continue to

focus on all profiles rather than a particular subset. It was noteworthy that the only provider type associated with a decreased likelihood for disagreement with the SME was family practice physician, which is contrary to the abovementioned expectations about aerospace medicine physicians.

Future attempts at modeling predictors of occupational limitations would benefit from using a larger dataset. Unfortunately, while electronic health records have allowed increased datafication of healthcare transactions, determination of the "true outcome" of whether a service member has occupational limitations ─ a necessary input to the modeling ─ still must rely on a SME reviewing the medical record, which is a time-intensive task. Meanwhile, continued training and quality assurance activities with feedback to PCPs to enable continuous learning are the most tractable near-term approach for improving profiling system performance.

6.0 REFERENCES

1. Agency for Healthcare Research and Quality. HCUP tools and software. 2017. [Accessed 22 Jan 2017]. Available from http://www.hcup-us.ahrq.gov/tools_software.jsp.
2. R Core Team. R: a language and environment for statistical computing. 2015. [Accessed 4 Mar 2017]. Available from http://www.R-project.org/.
3. van Belle G. Statistical rules of thumb, 2nd ed. Hoboken (NJ): Wiley-Interscience; 2008:125-126.
4. Chen T, He T. xgboost: eXtreme gradient boosting, R Package version 0.6-4. 2017. [Accessed 4 Mar 2017]. Available from http://cran.fhcrc.org/web/packages/xgboost/vignettes/xgboost.pdf.
5. Friedman JH. Stochastic gradient boosting. Comput Stat Data Anal. 2002; 38(4):367-378.
6. Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat. 2001; 29(5):1189-1232.

LIST OF ABBREVIATIONS AND ACRONYMS

AF	Air Force
AFSC	Air Force Specialty Code
AUROC	area under the receiver operating characteristic
CI	confidence interval
ICD-9-CM	International Classification of Diseases, Ninth Revision, Clinical Modification
IDMT	independent duty medical technician
MRI	magnetic resonance imaging
MTF	medical treatment facility
OR	odds ratio
PCP	primary care provider
PHA	Periodic Health Assessment
SE	standard error
SME	subject matter expert